

# 개요: 인도(人道) 보행 영상 데이터셋이란?

2017년 기준, 보건복지부 시,도 장애인등록현황에 따르면 보행에 어려움을 겪는 지체장애인 및 시각장애인은 약 150만 7천명 집계되었다. 기존의 문제 해결 방법으로 안내견, 흰지팡이 활용 및 보조기기(ETA, Electronic Travel Aids)가 활용되고 있으나 비용, 기술력, 사용성을 고려하지 않은 설계 등으로 한계를 가진다. 따라서 기존 방법의 한계를 극복하기 위해 장애인이 객체(이동체, 고정체), 객체간 거리, 노면 정보를 사용자 (시각 장애인, 지체장애인 등)에게 제공함으로써 보행 간 발생할 수 있는 위험을 회피하고 안전하고 편리한 이동을 보조하는 서비스 개발이 필요하다. 해외의 경우 인공지능 기술 중 가장 성숙도가 높다고 평가되는 컴퓨터 비전 기술을 활용해, 장애인 이동권 향상을 위한 다양한 서비스 개발이 이뤄지고 있으나 국내의 경우, 데이터셋의 부족이 제약 상황으로 작용하여 관련 인공지능 모델 및 서비스 개발이 활발히 이뤄지고 못하는 실정이다.

보행자용 인도는 차량용 도로와 다른 특성을 가지기 때문에 기존 데이터를 활용하여 보행자를 대상으로 한 인공지능 기술 및 응용 서비스 개발의 어려움이 존재한다. 따라서 본 인도 보행 영상 데이터셋은 장애인 인도 보행의 어려움, 장애인 이동권 문제와 같은 사회 문제 해결의 토대를 마련하고 인도 보행 인공지능 공개 데이터셋의 부족 문제를 보완하여 보행자를 위한 인공지능 기술 개발을 제고하기 위해 구축되었다.

## 데이터셋의 구성

본 데이터셋의 구성은 표 1. 과 같다. 제시된 데이터셋의 총 수량은 연구용으로도 충분한 양이며, 상용화 레벨에서도 강력한 응용 서비스 모델을 개발할 수 있는 양이다. 본 데이터셋을 활용하여 인도 보행자를 위한 객체/거리 인식 모델, 인도 노면 안전성 모니터링 시스템, 보행 가능 구역 인식 모델 등과 같은 서비스 모델 개발이 가능하다.

표 1. 데이터셋 구성 및 내용

데이터 가공 방식	수량	포함 내용
Bounding Box	35만	인도 보행에 장애가 되는 객체 29종에 대해 인공지능 학습 용으로 가공한 데이터
Polygon Segmentation	10만	인도 보행에 장애가 되는 객체 29종에 대해 인공지능 학습 용으로 가공한 데이터
Surface Masking	5만	인도 노면 상태 (재질, 특수성, 파손여부 등)정보를 폴리곤 형태로 어노테이션하고 마스킹하여 인공지능 학습용으로 가공한 데이터
Depth Prediction	15만	스테레오 카메라로 촬영한 인도 보행 이미지를 토대로 구축한 깊이 인식 데이터

# 데이터셋의 설계 기준과 분포

본 데이터셋을 구축할 때 장애인 인도보행을 위한 '이동권/접근권' 향상에 기여하기 위해 장애인협회 관계자와의

심층 인터뷰 및 자료 조사를 통해 장애인들의 인도보행에 있어서의 어려움과 문제점을 파악하여 이를 데이터셋 설계에 반영하였다. 특히, 인도 보행의 어려움으로서 인도상의 다양한 장애물, 노면의 재질, 파손, 특수성 등을 고려하여 이를 반영한 다양한 객체 클래스를 설계하였다.

**표 2. 장애물에 따른 객체 클래스**

분류	객체 클래스
이동체	<ul style="list-style-type: none"> <li>person (사람)</li> <li>car (승용차, 승합차)</li> <li>bus (버스)</li> <li>truck (트럭)</li> <li>bicycle (자전거)</li> <li>motorcycle (오토바이, 스쿠터)</li> <li>movable_signage (이동식 홍보 전시물/안내판)</li> <li>scooter (변속 기어가 없는 이륜차, 두 발로 탈 수 있는 기구)</li> <li>stroller (유모차)</li> <li>wheelchair (휠체어)</li> <li>dog (개)</li> <li>cat (고양이)</li> <li>carrier (리어카, 손수레)</li> </ul>
고정체	<ul style="list-style-type: none"> <li>tree_trunk (가로수 기둥)</li> <li>potted_plant (화분)</li> <li>traffic_light (신호등)</li> <li>traffic_sign (교통 표지판)</li> <li>pole (대/기둥)</li> <li>bench (2인 이상이 앉는 기구)</li> <li>chair (1인이 앉을 수 있는 기구)</li> <li>table (탁자)</li> <li>stop (버스/택시 정류장)</li> <li>kiosk (한쪽이 열린 점포)</li> <li>fire_hydrant (소화전)</li> <li>parking_meter (주차요금정산기)</li> <li>bollard (볼라드)</li> <li>barricade (바리케이드)</li> <li>power_controller (전력제어함)</li> <li>traffic_light_controller (신호등제어기)</li> </ul>

**표 3. 노면에 대한 객체 클래스 및 속성**

노면 클래스	속성
Sidewalk (인도)	<ul style="list-style-type: none"> <li>blocks (보도블럭)</li> <li>cement (시멘트)</li> <li>urethane (우레탄)</li> <li>asphalt (아스팔트)</li> <li>soil_stone (흙, 돌, 비포장도로)</li> <li>damaged (인도의 파손)</li> <li>other (정의된 인도의 속성 외의 경우)</li> </ul>
braille_guiding_block (점자블럭)	<ul style="list-style-type: none"> <li>normal</li> <li>damaged (점자블럭 파손)</li> </ul>
roadway (차도)	<ul style="list-style-type: none"> <li>normal</li> <li>crosswalk (횡단보도)</li> </ul>
alley (이면도로)	<ul style="list-style-type: none"> <li>normal</li> <li>crosswalk (횡단보도)</li> <li>speed_bump (과속방지턱)</li> <li>damaged (이면도로 파손)</li> </ul>
bike_lane (자전거도로)	(해당없음)
caution_zone (주의구역)	<ul style="list-style-type: none"> <li>stairs (계단)</li> <li>manhole (맨홀)</li> <li>tree_zone (가로수 한그루 식재를 위한 바닥면 구역)</li> <li>grating (그레이팅)</li> <li>repair_zone (보수구역)</li> </ul>

표 4. 데이터셋 구성 및 세부 데이터 분포

데이터 가공 방식	이미지 수	xml 수	디렉토리 수	디렉토리 1개당 구성 내용
Bounding Box	352,901	2,520	2,520	Bbox_xxxx\$image(.jpg or .png) 50~500장 + xml 1개
Polygon Segmentation	100,712	1,994	1,994	Polygon_xxxx\$image(.jpg or .png) 50~300장 + xml 1개
Surface Masking	51,318	900	900	Surface_xxx\$image (.jpg) 31~120장 + xml 1개 + 'MASK'하위폴더 \$mask image(.png) 31~120개
Depth Prediction	174,080 set (1세트 8개 이미지)	(해당없음)	93	Depth_xxx\$image(.jpg or .png) 8×n장 + calibration (.conf) 1개

## 데이터 구조

데이터셋에서 제공되는 어노테이션 정보를 담고 있는XML Data 에 대한 정보는 표 5. ~ 표 7. 과 같다.

표 5. meta 데이터

Tag		설명
task	id	task의 고유번호. 서버에서 관리하는 값으로 자동으로 부여
	name	task의 이름
	size	task의 사이즈
	mode	task의 작업 모드 (default : annotation) annotation은 이미지, interpolation은 동영상 작업
	overlap	task가 동영상인 경우 task간 겹쳐진 프레임의 개수
	bugtracker	이슈 관리
	flipped	뒤집힘 여부
	created	task가 생성된 시간
	updated	task가 수정된 시간
	start_frame	task의 시작 프레임 정보 (default : 0)
	stop_frame	task의 종료 프레임 정보 (default : 0)
	frame_filter	프레임 필터 정보 (default : 0)
	labels	라벨 리스트 (라벨은 label name, attribute로 구성)
	segments	task를 등록했을 경우 segments 관련 정보
owner	annotation 작업자 관련 정보	
dumped	xml 다운로드 받은 시간	

표 6. 어노테이션 정보

Tag	Attribute	설명
image	id	image 고유번호
	name	image 이름
	width	image 너비
	height	image 높이
box	label	레이블명
	occluded	가림/잘림 표시
	xtl/ytl	box에서 좌상단 x/y 좌표(X Top Left/Y Top Left)
	xbr/ybr	box에서 우하단 x/y 좌표(X Bottom Right/Y Bottom Right)
	z_order	객체가 그려진 순서 정보
polygon	label	레이블명
	points	polygon에서 폴리곤의 포인트 좌표
	occluded	가림/잘림 표시
	z_order	객체가 그려진 순서 정보
	group_id	polygon에서 그룹화 작업의 정, group_id 값이 같으면 같은 그룹
attribute	name	속성값

표 7. 깊이 추정 데이터셋의 경우 Depth\_xxx 폴더당 8×N 개의 파일과 calibration 파일 하나를 담고 있다.

표 7. 깊이 추정 데이터셋

Tag	Attribute	설명
image	id	image 고유번호
	name	image 이름
	width	image 너비
	height	image 높이
box	label	레이블명
	occluded	가림/잘림 표시
	xtl/ytl	box에서 좌상단 x/y 좌표(X Top Left/Y Top Left)
	xbr/ybr	box에서 우하단 x/y 좌표(X Bottom Right/Y Bottom Right)
	z_order	객체가 그려진 순서 정보
polygon	label	레이블명
	points	polygon에서 폴리곤의 포인트 좌표
	occluded	가림/잘림 표시
	z_order	객체가 그려진 순서 정보
	group_id	polygon에서 그룹화 작업의 정, group_id 값이 같으면 같은 그룹
attribute	name	속성값

# 데이터 예시

상기 '데이터 구조' 단락에서 설명한 xml 파일의 예시는 그림 1.과 같다. 또한 인도(人道) 보행 영상 데이터셋에 대한 DEMO 영상은 다음의 URL에서 시청이 가능하다.

- <https://www.youtube.com/watch?v=jQb3Rat6rxM> (도로 표면 마스크)
- <https://www.youtube.com/watch?v=F04eyIFGr3E> (객체인식)

```
<?xml version="1.0" encoding="utf-8" ?>
- <annotations>
  <version>1.1</version>
  - <meta>
    - <task>
      <id>5419</id>
      <name>0729_18(최종검수)</name>
      <size>396</size>
      <mode>annotation</mode>
      <overlap>0</overlap>
      <bugtracker />
      <flipped>False</flipped>
      <created>2019-11-25 10:02:02.210210+03:00</created>
      <updated>2019-11-25 11:09:09.085018+03:00</updated>
      <start_frame>0</start_frame>
      <stop_frame>0</stop_frame>
      <frame_filter>0</frame_filter>
    + <labels>
    + <segments>
    - <owner>
      <username>aiadmin</username>
      <email>aiadmin@testworks.co.kr</email>
    </owner>
    </task>
    <dumped>2019-11-27 09:39:09.251527+03:00</dumped>
  </meta>
  - <image id="0" name="MP_KSC_012801.jpg" width="1920" height="1080">
    <box label="movable_signage" occluded="0" xtl="432.25" ytl="509.28" xbr="659.80" ybr="1080.00" z_order="7" />
    <box label="movable_signage" occluded="0" xtl="906.00" ytl="554.02" xbr="968.25" ybr="721.48" z_order="3" />
    <box label="movable_signage" occluded="1" xtl="1159.20" ytl="572.40" xbr="1193.74" ybr="652.09" z_order="2" />
    <box label="tree_trunk" occluded="1" xtl="1674.91" ytl="300.86" xbr="1724.25" ybr="698.90" z_order="1" />
    <box label="traffic_light" occluded="1" xtl="1051.89" ytl="444.31" xbr="1084.27" ybr="455.28" z_order="9" />
  - <image id="4" name="MP_SEL_PN001105.jpg" width="1920" height="1080">
    <polygon label="person" occluded="0"
      points="1364.70,725.20;1356.10,708.40;1357.10,690.00;1356.80,657.90;1364.40,619.79;1367.14,583.24;1361.66,523.
      z_order="7" />
    <polygon label="car" occluded="0"
      points="1920.00,714.43;1880.10,724.96;1864.31,631.54;1859.04,565.75;1862.99,465.75;1916.94,461.80;1918.25,519.
      z_order="1" group_id="10" />
    <polygon label="person" occluded="0"
      points="1128.65,714.82;1135.05,686.49;1135.96,647.20;1133.22,585.06;1126.82,557.65;1132.31,529.33;1136.88,536.
      z_order="8" />
    <polygon label="car" occluded="0"
      points="1715.62,534.17;1699.83,561.80;1676.15,565.75;1668.25,571.01;1677.46,582.85;1655.09,614.43;1647.20,642.
      z_order="3" group_id="10" />
  - <image id="6" name="MP_SEL_SUR_051608.jpg" width="1920" height="1080">
    - <polygon label="sidewalk" occluded="0"
      points="904.31,1080.00;884.31,0.00;283.77,0.00;285.66,23.73;265.40,28.82;262.23,19.33;195.78,56.03;222.37,180.0
      z_order="2">
      <attribute name="attribute">blocks</attribute>
    </polygon>
    - <polygon label="caution_zone" occluded="0"
      points="222.29,180.32;480.35,182.23;363.84,356.36;0.00,359.94;0.00,186.45;20.70,176.70;41.46,176.68;131.70,177.
      z_order="3">
      <attribute name="attribute">tree_zone</attribute>
    </polygon>
    - <polygon label="sidewalk" occluded="0" points="119.56,109.22;131.33,177.26;64.61,173.98;58.06,149.78"
      z_order="4">
      <attribute name="attribute">blocks</attribute>
    </polygon>
    <polygon label="bike_lane" occluded="0"
      points="905.21,1079.63;884.43,0.00;1188.10,0.00;1370.45,242.31;1543.73,432.60;1920.00,850.79;1920.00,1080.00"
      z_order="1" />
  </image>
```

그림 1. 어노테이션 xml 파일 예시

# 데이터 구축 과정

본 데이터셋의 구축 과정은 표 8. 과 같다.

표 8. 인도(人道) 보행 영상 데이터셋 구축 과정

단계	세부 절차	설명
1. 설계	1.1 데이터 설계	장애인협회와의 인터뷰 및 자료 조사를 기반으로 다양한 객체 클래스 설계 및 데이터 구조 설계
2. 수집	2.1 원천 데이터 선정	원천 데이터 수집 방법 및 규모 설정
	2.2 원천 데이터 수집 교육	데이터 수집 및 정제 기준 교육
	2.3 원천 데이터 수집	모바일폰과 ZED Stereo 카메라로 촬영
3. 정제	3.1 원천 데이터 검수	영상의 흔들림, 촬영 상황 등을 지정된 숙련 검수자가 검수
	3.2 원천 데이터 비식별	수집된 영상 데이터에 의도하지 않게 포함된 개인 정보 비식별화
4. 가공	4.1 가공 인력 양성 교육 시행	인공지능 학습용 데이터 가공에 필요한 작업 교육과 훈련
	4.2 데이터 가공	<ul style="list-style-type: none"> <li>• 바운딩 박스(Bounding Box)</li> <li>• 폴리곤 세그멘테이션(Polygon Segmentatio</li> <li>• 서피스 마스킹(Surface Masking)</li> <li>• Depth Prediction</li> </ul>
5. 검수	데이터 검수	<ul style="list-style-type: none"> <li>• 1단계 : 작업자 간 검수</li> <li>• 2단계 : 매니저 검수</li> <li>• 3단계 : 최종 검수</li> </ul>

## 검수와 품질 확보

인도(人道) 보행 영상 데이터셋 품질은 아래와 같이 정의한다.

$$(\text{데이터 품질}) = \frac{x - (y_1 + y_2 + y_3 + y_4)}{x} \times 100$$

$x$  = 이미지 내 어노테이션해야 할 객체의 수  
 $y_1$  = 오검출된 객체의 수  
 $y_2$  = 미검출된 객체의 수  
 $y_3$  = 과검출된 객체의 수  
 $y_4$  = 어노테이션 범위 불량인 객체의 수

인도보행 영상 데이터셋 가이드라인을 기준으로 어노테이션 작업을 수행하여 가공된 데이터셋의 '데이터 품질'은 검수를 통해 얻은 오검출, 미검출, 과검출, 어노테이션 범위 불량인 객체의 빈도를 위의 식으로 산출하여 결정한다. 해당 데이터셋에 오검출, 미검출, 과검출, 어노테이션 범위 불량인 객체의 빈도가 낮을수록 데이터 품질 지표가 높아진다. 바운딩박스과 뎀스프리딕션은 인도보행 영상 데이터셋의 품질 정의를 동일하게 적용한다.

폴리곤 어노테이션은 인도보행 영상 데이터셋 품질 정의를 토대로 아래와 같이 변경하여 데이터 품질을 정의한다.

$$(\text{데이터품질}) = \frac{x - (y_1 + y_2 + y_3 + y_4 + y_5)}{x} \times 100$$

- $x$  = 이미지 내 어노테이션해야 할 객체의 수
- $y_1$  = 오검출된 객체의 수
- $y_2$  = 미검출된 객체의 수
- $y_3$  = 과검출된 객체의 수
- $y_4$  = 어노테이션 범위 불량인 객체의 수
- $y_5$  = 객체 간 그룹화 오류의 수

서피스 마스킹은 인도보행 영상 데이터셋 품질 정의를 토대로 아래와 같이 변경하여 데이터 품질을 정의한다.

$$(\text{데이터품질}) = \frac{x - (y_1 + y_2 + y_3 + y_4 + y_5)}{x} \times 100$$

- $x$  = 이미지 내 어노테이션해야 할 객체의 수
- $y_1$  = 오검출된 객체의 수
- $y_2$  = 미검출된 객체의 수
- $y_3$  = 과검출된 객체의 수
- $y_4$  = 어노테이션 범위 불량인 객체의 수
- $y_5$  = 어노테이션 순서 불량인 객체의 수

본 데이터셋은 아래와 같은 3단계의 검수 절차를 거쳐 오검출, 미검출, 과검출, 어노테이션 범위 불량인 객체가 없는지 검수하여 데이터 품질을 관리한다.

- Peer Review : 클라우드 소싱으로 참여한 작업자 중 높은 수행 성과를 달성한 작업자 간 리뷰 진행, 데이터 품질 95% 확보
- Manager Review : 작업자와 프로젝트 매니저 간에 리뷰 진행, 품질 기준에 미달한 데이터셋은 재작업 수행으로 품질 98% 확보
- Final Review : 최종 전달 이전에 2단계 검수 절차에서 품질 기준을 충족한 데이터 중 샘플 데이터 추출 및 검수를 통해 데이터 품질 99% 확보



그림 7 품질 확보를 위한 3단계 품질 검수 체계

## 데이터 구축 담당자

주식회사 테스트웍스 AI사업본부 : [ai@testworks.co.kr](mailto:ai@testworks.co.kr)